Open Set Radar HRRP Recognition Based on Random Forest and Extreme Value Theory

Yanhua Wang*, Wei Chen*, Jia Song*, Yang Li*‡, Xiaopeng Yang*
*School of Information and Electronics, Beijing Institute of Technology, Beijing, China
*Beijing Key Laboratory of Embedded Real-time Information Processing Technology,
Beijing Institute of Technology, Beijing, China
‡Email: bit_liyang@bit.edu.cn

***Abstract*—Most of the progresses achieved in radar high range resolution profile (HRRP) recognition rely on the closed set condition, where the test sample is from a known class. In realistic scenario, however, the test sample may be drawn from unknown classes, which is regarded as an open set recognition task. In such cases, conventional recognition algorithms will inevitably make a wrong prediction. In this paper, the open set problem is addressed by incorporating the extreme value theory (EVT) into the random forest (RF) classifier. The outputs of RF are analyzed to determine whether the test sample should be rejected as an unknown class. At the training phase, a Weibull-based extreme-value meta-recognition is introduced to describe the statistical characteristics of the known classes. At the testing phase, a probability estimation method is introduced to compute the probabilities of the test sample belonging to known and unknown classes based on trained Weibull distributions. The test sample is assigned to the class of highest probability. Experimental results demonstrate that the proposed method outperforms the state-of-art NN, 1-vs-set machine and W-SVM in rejecting unknown classes.**

Keywords*—*open set recognition, random forest, high range resolution profile (HRRP), extreme value theory.

I. Introduction

The common approaches in radar automatic target recognition (RATR) are based on high range resolution profiles (HRRP) [1], synthetic aperture radar (SAR) images [2] and inverse synthetic aperture radar (ISAR) images [3]. HRRP represents the one-dimensional (1D) distribution of an object's scattering centers along the radar line of sight, while SAR/ISAR images show the two-dimensional (2D) distribution in the slant-range plane. Due to its mild complexity in acquisition and processing, HRRP plays an important role in RATR community [4]. Most of the existing HRRP recognition algorithms are designed for a closed set world, whereby all testing classes are assumed to be known at the training phase. When testing a sample from unknown classes, these algorithms will assign it to a known label, which is unacceptable. This problem is known as open set recognition, which aims to reject the sample from unknown classes and put the known samples into correct classes at the same time. Table I shows the differences between closed set and open set recognition.

This work was supported by the National Natural Science Foundation of China (Grant No. 61701026), Chang Jiang Scholars Programme (Grant No. T2012122) and 111 Project of China (Grant No. B14010).

TABLE I: Differences between closed set and open set. Here N, M denote the number of classes in different phases and $2 \leq N < M$.

	Classifier	**Training**	**Target**	**Testing**
Open set	Binary	2	2	M
	Multi-class	N	N	M
Closed set	Binary	2	2	2
	Multi-class	N	N	N

Open set recognition is challenging due to the incomplete knowledge at the training stage. A few methods have been proposed in literature. A first category is based on nearest neighborhood (NN) with an additional threshold, but the estimation of a suitable threshold is difficult. As another approach, Scheirer extended support vector machine (SVM) to the open set world and proposed 1-vs-set [5] and W-SVM [6] for binary and multi-class cases, respectively. In 1-vs-set machine, an additional hyperplane is introduced to separate the known and unknown classes and it is determined by the greedy searching strategy. W-SVM involves two separate SVMs to yield the probabilities of test sample unrelated and belonging to known classes, respectively. The two probabilities are fused to determine whether to reject a test sample. The above SVM-based open set recognition methods have been applied in classification tasks such as SAR [7], laser detecting and ranging (LADAR) [8] and forward-looking infrared radar (FLIR) [9]. Recently, a sparse representation based open-set recognition method is proposed in [10]. The tail distributions of matched and non-matched reconstruction errors are fused to jointly determine the identity of the test sample. With the development of deep learning, an open set deep network is proposed to offer an end-to-end capability for open set recognition [11].

In this paper, we propose an open set recognition framework based on RF and apply it to HRRP recognition. The outputs of RF are regarded as the 'high-level features' of the input sample, which are used to estimate the distribution of the training data via extreme value theory (EVT) [12]. At the training phase, the Weibull-based extreme value distribution is employed to establish a 'boundary' for the known classes because it can predict the probability of a rare value occurring. An extreme-value meta-recognition algorithm is applied to fit the Weibull distributions. At the testing phase, the probabilities of the test sample belonging to known and unknown classes are calculated based on the above Weibull distribution. The test sample is assigned to the class of the highest probability.

The structure of this paper is organized as follows: The structure of open set RF and its principle are described in Section 2. Section 3 demonstrates the experimental results and the conclusion is given in Section 4.

II. Open set random forest

A. Structure

RF is an ensemble classifier that is made up of many decision trees (DTs) [13]. Fig. 1(a) shows the structure of RF for closed set world. The outputs of all DTs are combined together to generate the final prediction result. RF is widely applied because of its high efficiency and robustness to data-lacking problems. To deal with the open set problem, we add an open set module to analyze the RF outputs. Fig. 1(b) shows the structure of open set RF.

Fig. 2 illustrates the flowchart of the open set module, which consists of two parts. In the training part, we apply an extreme-value meta-recognition method to fit Weibull distributions to known data. In the testing part, a probability estimation method is used to compute probabilities of the test sample belonging to all classes including unknown classes based on the similarities between the test sample and each known class. The label of the test sample is then assigned according to the highest probability and the test sample is rejected when it belongs to unknown classes.

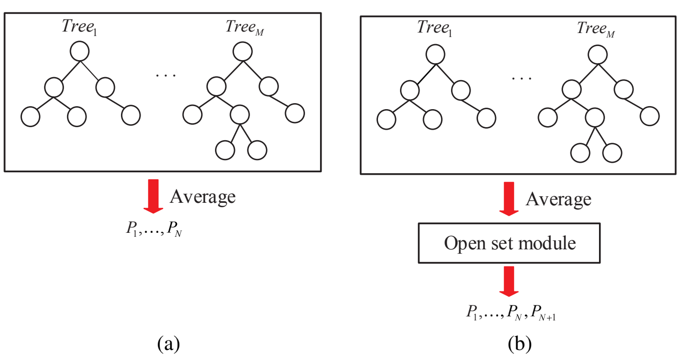

Fig. 1: The structure of RF: (a) closed set RF; (b) open set RF

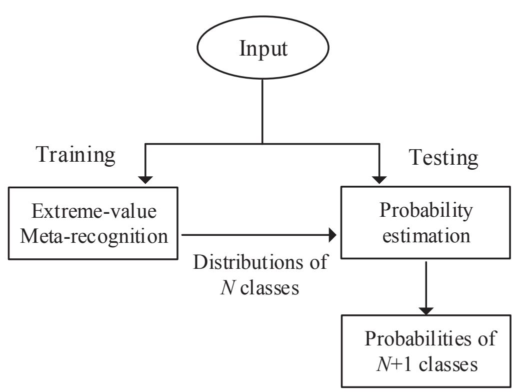

Fig. 2: The flowchart of open set module.

Algorithm 1 Extreme-value meta-recognition algorithm.

Require: The outputs vector of RF
Output: Distribution of each known classes $W_i(\lambda_i, \kappa_i, \boldsymbol{\beta}_i)$
for $i = 1, \ldots N$ **do**
  Compute the class mean vector;
  Compute the distance between samples and class mean vector;
  Consider top m samples as extreme value point based on distance;
  Fit Weibull distribution $W_i(\lambda_i, \kappa_i, \boldsymbol{\beta}_i)$ based on extreme value point;
end for

B. Extreme-value Meta-recognition

In open set recognition, the most important issue is to determine the boundary between known and unknown classes. The difficulties lie on the lack of knowledge of unknown classes. In this paper, we use probabilities to determine known and unknown classes rather than distance. Here, the extreme-value meta-recognition is used to determine boundary between the known and unknown classes automatically. In general, meta-recognition is defined as a performance prediction technique with the capability of observing and even adjusting the recognition process [14]. EVT describes the distribution of data of abnormally high or low values [12]. Thus, extreme-value meta-recognition offers an approach to evaluate the possibility that a test sample belongs to an unknown class. EVT states that the underlying extreme value distribution abides by one of three functions. The commonly used function corresponds to a Weibull distribution and it is given by

$$W(t) = \frac{1}{\lambda} e^{-v^{-1/\kappa}} v^{-(1/\kappa+1)}, \tag{1}$$

where $v = 1 + \kappa \frac{t-\beta}{\lambda}$ where β, λ, κ are the location, scale, and shape parameters respectively.

The extreme-value meta-recognition algorithm is described in Algorithm 1. Firstly, the outputs of RF are used to calculated the mean vector of each class. Then, the training samples are sorted in descending order according their Euclidean distances to the mean vector. The top m samples are considered as extreme value point for each class. Finally, the training data of each class are used to fit Weibull distributions based on distance.

C. Probability Estimation

The goal of probability estimation is obtaining probabilities of the test sample belonging to each class. The primary work is computing the rejection probability, which refers to the probability of the sample belonging to unknown classes. However, it is difficult to directly compute the rejection probability without prior information. Here, we first calculate the probabilities of the test sample belonging to known classes using the Weibull cumulative distribution function (CDF) with parameters estimated above [14]. Then, the rejection probability is indirectly calculated through these probabilities. The probability estimation algorithm is summarized as Algorithm 2.

For the RF output $\mathbf{x}$ of a test sample, we derive the probability of the test sample belonging to an known distribution as follow:

$$S\left(y|\mathbf{x}\right) = 1 - e^{-\left(\frac{|\mathbf{x}-\boldsymbol{\beta}|}{\lambda}\right)^{\kappa}}, \tag{2}$$

The output of RF is the per-class estimation score, therefore combining the outputs with Weibull CDF to estimate the similarity is more accurate, as follow:

$$S\left(C_i|\mathbf{x}\right) = \mathbf{x}\left(i\right) \cdot \left(1 - e^{-\left(\frac{|\mathbf{x}-\boldsymbol{\beta}_i|}{\lambda_i}\right)^{\kappa_i}}\right), \tag{3}$$

where C_i is the known class i in the training. Meanwhile, we consider top ranks provide a more meaningful probability for estimations of known classes in open set RF. The probability equation can be revised by using a weight value as (4).

$$S\left(C_i|\mathbf{x}\right) = \mathbf{x}\left(i\right) \cdot \left(1 - \frac{N-j}{N} e^{-\left(\frac{|\mathbf{x}-\boldsymbol{\beta}_i|}{\lambda_i}\right)^{\kappa_i}}\right), \tag{4}$$

where j is the rank of class i after sorting based on the outputs of RF.

After getting the probabilities of the test sample belonging to known classes, the key work is computing the rejection probability. The rejection probability for the test sample is derived based on reverse Weibull CDF, which defined as:

$$S\left(C_{N+1}|\mathbf{x}\right) = \sum \mathbf{x}\left(i\right) \cdot \frac{N-j}{N} e^{-\left(\frac{|\mathbf{x}-\boldsymbol{\beta}_i|}{\lambda_i}\right)^{\kappa_i}}, \tag{5}$$

where index $N+1$ denotes unknown classes. It is necessary to normalize probabilities because the sum of probabilities is 1. The normalization formula is defined as (6).

$$P_i = \frac{e^{S(C_i|\mathbf{x})}}{\sum_{k=1}^{N+1} e^{S(C_k|\mathbf{x})}}. \tag{6}$$

Finally, the sample is assigned to the class of highest probability based on computed probabilities. If the test sample is assigned to unknown classes, it is rejected to avoid reducing the classifier recognition performance.

Algorithm 2 Probability estimation algorithm.

Require: Weibull distribution $W_i(\lambda_i, \kappa_i, \boldsymbol{\beta}_i)$
Input: The RF outputs of the test sample
Output: Probabilities of the test sample belonging to each class

for $i = 1$, ..., N **do**
  Compute the probabilities of the test sample belonging to known class i ;
end for
Compute the rejection probability ;
Probability normalization;

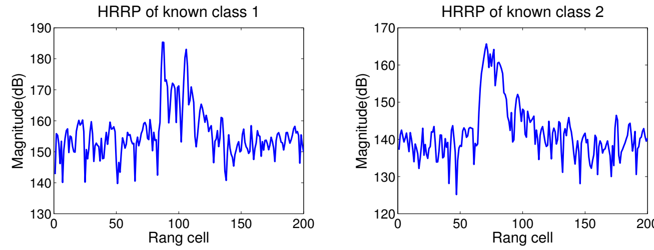

Fig. 3: The HRRPs of two known classes.

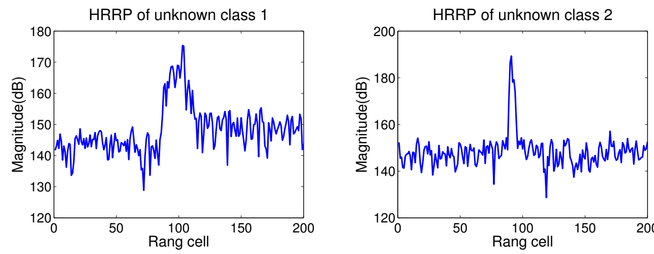

Fig. 4: The HRRPs of two unknown classes.

III. Experimental Results and Analysis

A. Experiment Setup

An experimental radar with a bandwidth of 1200MHz was developed to collect the HRRP data. Each HRRP sample contains of 200 points. In our test, four targets were used as known classes and others were regarded as unknowns. At the training phase, each kind of target contains 800 samples, while at testing phase each kind of target contains 500 samples. Typical HRRPs of the known and unknown targets are shown in Fig. 3 and Fig. 4, respectively.

Before classification, segmentation and feature extraction were performed. The aim of segmentation was to determine the range intervals of the target and a two-threshold detection algorithm was applied [15]. Then, twelve features were extracted, including length, number of strong scattering points, sparsity of scattering points, standard deviation, etc. In our experiment, RF consists of 40 DTs and the max depth of a DT is set as 10. To make full use of extracted features, we set the maximum number of selected features as 12 for each DT.

Generally, a binary confusion matrix about known and unknown classes is used to describe the experimental results of open set recognition, as shown in Table II.

TABLE II: The confusion matrix of known and unknown classes

	known classes	unknown classes
known classes	TP	FP
unknown classes	FN	TN

In Table II, TP is the number of true known samples, FN is the number of false unknown samples, FP is the number of false known samples, TN is the number of true unknown samples.

In open set recognition, recognition accuracy rate (Acc) and F1-measure (F1) are usually used to evaluate the performance of different recognition algorithms. The higher Acc and F1, the better performance. Acc is defined as

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TN} + \mathrm{FN} + \mathrm{FP} + \mathrm{TP}}. \tag{7}$$

F1 is a harmonic mean of Precision and Recall. The higher F1 the better performance of a recognition algorithm and F1 is defined as

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{8}$$

where Precision and Recall are defined as follow:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{9}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{10}$$

B. Results

The recognition performance of the proposed method is evaluated by comparing with three other open set recognition algorithms NN, 1-vs-set and W-SVM. Experimental results of different algorithms are given in Table III. It can be seen that when the number of unknown classes is less 3, the 1-vs-set has the highest Acc and F1. Meanwhile, the Acc of three other algorithms open set RF, NN and W-SVM is more than 90% and F1 is more than 0.94. With the number of unknown classes increasing, the performance of NN, 1-vs-set and W-SVM reduce rapidly while the proposed method stably remains a good performance. When having 3 unknown classes or more, the proposed method has highest Acc and F1. The performance of 1-vs-set reduces because second hyperplane determining by training data is impossible to reject all unknown classes. The yielded two probabilities of W-SVM based on known data is difficult to reject all unknown classes. While the performance of NN reduces because of the unsuitable threshold. The probability of test sample belonging unknown classes is computed based on probability of test sample belonging to each known class in open set RF. In our test, the known classes remain the same, thus open set RF stably remains a good performance when the number of unknown classes increasing.

TABLE III: The recognition performance of different algorithms

	Methods	1 unknown class	2 unknown classes	3 unknown classes	4 unknown classes
Acc	open set RF	90.92%	92.40%	**92.66%**	**93.37%**
	NN	94.76%	95.63%	90.40%	91.60%
	1-vs-set	**96.84%**	**97.23%**	91.60%	83.15%
	W-SVM	92.24%	93.53%	89.29%	90.63%
F1	open set RF	0.9423	0.9423	**0.9354**	**0.9331**
	NN	0.9682	0.9682	0.9223	0.9223
	1-vs-set	**0.9806**	**0.9796**	0.9313	0.8557
	W-SVM	0.9501	0.9501	0.9078	0.9078

IV. Conclusion

In this paper, a novel open set HRRP recognition framework based on RF is proposed. An open set module is introduced to adapt RF to open set recognition. At the training phase, the outputs of RF are used to fit Weibull distributions of known classes using an extreme-value meta-recognition algorithm. At the testing phase, a probability estimation is introduced to compute probabilities of the test sample belonging to known classes based on trained Weibull distributions. The rejection probability is indirectly computed based on these probabilities. The test sample is assigned to the class of highest probability and it is rejected when the rejection probability is highest. Experimental results demonstrate that the proposed open set RF can automatically reject many unknown classes without any prior information of unknown classes. When unknown classes increase, open set RF has a better performance than NN, 1-vs-set and W-SVM. Meanwhile, the proposed method has a more stable recognition performance comparing with NN, 1-vs-set and W-SVM.

References

[1] S. P. Jacobs, and J. A. Osullivan, "Automatic target recognition using sequences of high resolution radar range-profiles," *IEEE Transactions on Aerospace and Electronic Systems.*, vol. 36, pp. 364–381, Jan. 1999.

[2] C. Tison, N. Pourthie, and J. C. Souyris, "Target recognition in SAR images with support vector machines (SVM)," *IEEE International Geoscience. Remote Sensing Symposium*, 2007, pp. 456–459.

[3] W. Ning, W. Chen, and X. Zhang, "Automatic target recognition of ISAR object images based on neural network," *IEEE International Conference on Neural Networks and Signal Processing*, 2003, pp. 373–376.

[4] L. Du, H. Liu, and Z. Bao, "Radar HRRP statistical recognition: parametric model and model selection," *IEEE Transactions on Signal Processing.*, vol. 56, pp. 1931–1944, May. 2008.

[5] W. J. Scheirer, A. Rocha, A. Sapkota, and T. E. Boult, "Towards open set recognition," *IEEE Transactions on Software Engineering.*, vol. 35, pp. 1757–1772, Jul. 2014.

[6] W. J. Scheirer, L. P. Jain, and T. E. Boult, "Probability models for open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 36, pp. 2317–2324, Nov. 2014.

[7] M. Scherreik, and B. Rigling, "Multi-class open set recognition for SAR imagery," *Proceedings of SPIE*, 2016, vol. 9844, pp. 98440M_1–98440M_9.

[8] M. Scherreik, and B. Rigling, "Ladar ATR via probabilistic open set techniques," *Proceedings of SPIE*, 2014, vol. 9090, pp. 90900A_1–90900A_9.

[9] M. Scherreik, and B. Rigling, "An evaluation of open set recognition for FLIR images," *Proceedings of SPIE*, 2015, vol. 9476, pp. 94760N_1–94760N_10.

[10] H. Zhang, and V. M. Patel, "Sparse representation-based open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 39, pp. 1690–1696, Aug. 2017.

[11] A. Bendale, and T. E. Boult, "Towards open set deep networks," *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1563–1572.

[12] W. J. Scheirer, A. Rocha, R. Micheals, and T. E. Boult, "Robust fusion: Extreme value theory for recognition score normalization," *European Conference on Computer Vision*, 2010, vol. 76, pp. 481–495.

[13] L. Breiman, "Random forest," *Machine Learning.*, vol. 45, pp. 5–32, Jan. 2001.

[14] W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boult, "Meta-recognition: The theory and practice of recognition score analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 33, pp. 1689–1695, Aug. 2011.

[15] T. Long, L. Zheng, Y. Li, and X. Yang, "Improved double threshold detector for spatially distributed target," *IEICE Transactions on Communications.*, vol. E95. B, pp. 1475–1478, Apr. 2012.